Encrypted Cartographic Masking for Geospatial Privacy Protection in Location Based Services

Keith C. Clarke
Department of Geography
University of California, Santa Barbara
11S KU38488 11890
Point Set Masking Methods

• Aggregation
• Movement to grid centers
• Move to street/tract center
• Anonymization (but reverse geocoding)
• Affine transformation
• Random or localized displacement
New method: Multiscale digit switching
MGRS

- Initial UTM zone number
- Grid cell designator
- Cell square identifier
- Easting 0-99999
- Northing 0-99999
Step 1: Use NGA’s GEOTRANS

- Original point set in lat, long + datum
- Put into GEOTRANS .DAT format
- Use GEOTRANS to convert to MGRS
2. Use custom program “hide”

- 10 digits to replace
- 1.3M possible combinations that assure all digits are altered
- Each combination is tried, but with replacement only for specific digit places (Brute Force)
- For each coding, compute set of metrics for new point set, compare with original
- Includes extremes, means, medians, standard deviations in $x$ and $y$
- Multiply difference sum by difference in nearest neighbor statistic
- Retain code with the smallest difference from the original point set’s statistics
For example

<table>
<thead>
<tr>
<th>Original Point Encoded</th>
<th>Level 3</th>
<th>MGRS coordinate</th>
<th>UTM Zone Number</th>
<th>Grid Cell Designator</th>
<th>Cell Square Identifier</th>
<th>Easting</th>
<th>Northing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point</td>
<td>10</td>
<td>10SEG4079590435</td>
<td>10</td>
<td>S</td>
<td>EG</td>
<td>40795</td>
<td>90435</td>
</tr>
<tr>
<td>Code key</td>
<td>NA</td>
<td>10SEG4031490684</td>
<td>10</td>
<td>S</td>
<td>EG</td>
<td>40314</td>
<td>90684</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10SEG9758640321</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

- 0 maps to 9
- 1 maps to 7
- 2 maps to 5
- 3 maps to 8
- 4 maps to 6
- 5 maps to 4
- 6 maps to 0
- 7 maps to 3
- 8 maps to 2
- 9 maps to 1

<table>
<thead>
<tr>
<th>Code last 3 digits</th>
<th>795</th>
<th>435</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 maps to 3</td>
<td>395</td>
<td>435</td>
</tr>
<tr>
<td>9 maps to 1</td>
<td>315</td>
<td>435</td>
</tr>
<tr>
<td>5 maps to 2</td>
<td>312</td>
<td>435</td>
</tr>
<tr>
<td>4 maps to 5</td>
<td>312</td>
<td>535</td>
</tr>
<tr>
<td>3 maps to 7</td>
<td>312</td>
<td>572</td>
</tr>
<tr>
<td>2 maps to 8</td>
<td>312</td>
<td>578</td>
</tr>
</tbody>
</table>

Complete: 40312 90578

MGRS Encoded: 10SEG4031490684
GPS trace: Golden Gate National Recreation Area, California
Applied to last two digits
Level 2 detail
Impact of scale/resolution

1

2

3

4

5
Conclusion

• Masking can be customized by scale
• Displacements from minimum to complete masking
• Spatial statistics for whole distribution preserved
• Code “reveal” takes stored code and inverts exactly
• Data + code contains all information, and aggregate statistics match
• Equivalent to lock and key for point data
Availability

- Code released to public domain on SourceForge multiscalemaskingofpoints.sourceforge.net
- Geotrans: earth-info.nga.mil/GandG/geotrans